

# Principally polarized abelian surfaces with surjective galois representations on $l$ -torsion

Erik Wallace

December 18, 2012

## 1 Introduction

A well known result of Duke [?] states that if  $\mathcal{C}(X)$  denotes the set of equivalence classes elliptic curves over  $\mathbb{Q}$  of the form  $y^2 = x^3 + rx + s$  such that  $\max\{|r|^3, |s|^2\} \leq X^6$ , and  $\mathcal{E}(X)$  denotes the subset for which the representation of  $\text{Gal}(\overline{\mathbb{Q}}/\mathbb{Q})$  on the  $l$ -torsion of a representative elliptic curve is not surjective, then

$$\lim_{X \rightarrow \infty} \frac{|\mathcal{E}(X)|}{|\mathcal{C}(X)|} = 0.$$

Cojocaru and Hall [?] have proven a generalization of this to one parameter families of elliptic curves, and Zywinia [?] has proven a generalization to arbitrary number fields. So far, however, there have been no extensions to higher dimensional abelian varieties. In this paper we prove a generalization to abelian surfaces.

Cojocaru and Hall [?] consider a curve  $C$  over  $\mathbb{Q}$ , and look at an elliptic curve  $E$  defined over the function field of  $C$ . Then they construct coverings  $C_l \rightarrow C$  that encapsulate the information about the  $l$ -level structure, and look at the Galois action on these coverings. If  $t$  is a rational point of  $C$ , and  $E_t$  denotes the corresponding elliptic curve, then we get a Galois representation on the  $l$ -torsion of  $E_t$ . Using this, they show that by bounding the height of  $t$ , a similar result to Duke's can be obtained. One shortcoming of their method is that they use the Riemann-Hurwitz formula, which is specific to dimension 1. However, we show that it is not necessary to use this. In our general setup we use a projective variety of arbitrary dimension  $r$ , which we specialize in our application to abelian surfaces by assuming  $r = 3$ , but otherwise it is similar.

Zywinia [?] generalizes Duke's original result to arbitrary number fields. This is accomplished with a version of the large sieve proven by Serre [?], which is applied for each  $l$ , and each conjugacy class in  $\text{GL}_2(\mathbb{F}_l)$ . Another interesting observation of Zywinia is that the Siegel-Walfisz theorem is not actually necessary. However, Zywinia does not consider elliptic curves over the function field of a curve, and so he is forced to use a result of Jones [?] that requires the Eichler-Selberg trace formula. We avoid this by using a version of the Chebotarev

density theorem for function fields first proven by Lang [?], which we sharpen slightly. It is worth noting that Cojocaru and Hall use a similar theorem, and that when it is used in conjunction with the large sieve, we get a result that is an effective form of Hilbert irreducibility.

In section 2 we begin discussing a version of the large sieve compatible with heights, as well as an application of the lower bound sieve that we need. A large portion of our results are not restricted by dimension in any way, and can possibly be applied to cases other than abelian varieties. So in section 3 we discuss the results that can be proven in a very general way, the main result being Theorem 6. This theorem is applied in section 4 to abelian varieties of dimension 1 and 2. The result in dimension 2 is new, but in dimension 1 it is little more than a hybrid of previous results. A list of notation can be found in appendix A

## 1.1 Acknowledgments

This paper comprises a portion of my dissertation. I would like to thank my advisor Michael Larsen for his guidance, and Aner Shalev for directing us to the work of Kleidman and Liebeck.

## 2 Sieve theory

The sieve methods we will use need to be compatible with heights on number fields. This is accomplished using the following construction which can be compared with the method of Schanuel [?]. Let  $K$  be a number field of degree  $d$  and ring of integers  $\mathcal{O}_K$ , and let  $S_\infty$  denote the set of infinite places of  $K$ . If  $v \in S_\infty$ , then

$$\|\cdot\|_v = |\cdot|_v^{[K_v:\mathbb{R}]/d}$$

defines a norm on  $K_v$ . If  $u = (u_0 : u_1 : \cdots : u_r) \in \mathbb{P}_K^r(K)$ , and  $\mathfrak{a}_u$  denotes the fractional ideal generated by the  $u_i$ , then we have

$$H(u) = N(\mathfrak{a}_u)^{-\frac{1}{d}} \prod_{v \in S_\infty} \sup_i \|u_i\|_v$$

where  $H$  is the absolute height. By scaling, it is possible to obtain coordinates  $u_i$  in  $\mathcal{O}_K$ . Making this choice minimally gives us a unique representative in a fundamental domain of  $K^{r+1} - 0^{r+1}$  under the action of units. Let  $\Lambda$  be the image of  $\mathcal{O}_K$  under the diagonal embedding  $K \rightarrow \prod_{v \in S_\infty} K_v$ , and let

$$B_K(x) = \{u \in \mathbb{P}_K^r(K) : H(u) \leq x\}.$$

With the coordinates of each  $u$  chosen as above, we first lift to  $K^{r+1} - 0^{r+1}$  and then consider the image in  $\prod_{v \in S_\infty} K_v^{r+1}$  under the diagonal embedding, as illustrated by the following diagram

$$\mathbb{P}_K^r(K) \longleftarrow K^{r+1} - 0^{r+1} \longrightarrow \prod_{v \in S_\infty} K_v^{r+1}.$$

It is contained in  $\Lambda^{r+1}$ , and by the results of Schanuel its size grows at a rate proportional to that of a ball of radius  $x$  in  $\Lambda^{r+1}$ . This will allow us to apply Serre's version of the large sieve (see [?] or [?]).

In [?] Kowalski has given a language for sieves that we would like to generalize slightly, and also use to discuss the constructions in our particular application. Then we will prove a version of the large sieve that is compatible with heights on affine space, and an application of the lower bound sieve that we will use with it.

A *sieve setting* is a triple  $(Y, \mathcal{A}, (\pi_\alpha))$  consisting of a set  $Y$ , an indexing set  $\mathcal{A}$ , and for each  $\alpha \in \mathcal{A}$  a map  $\pi_\alpha : Y \rightarrow Y_\alpha$ , where  $Y_\alpha$  is a finite set. For both the large sieve and the lower bound sieve, we will take  $Y = \Lambda^{r+1}$ . The indexing set  $\mathcal{A}$  will be different in each case though: for the large sieve we will take it to be  $\Sigma_K$ , the set of non-zero prime ideals in  $\mathcal{O}_K$ , whereas for the lower bound sieve we will use the ordinary primes, which we will view as ideals in  $\mathcal{O}_K$ . If  $\mathfrak{a}$  is an arbitrary ideal in  $\mathcal{O}_K$ , then it can be identified with a sublattice of  $\Lambda$ , and the quotient is isomorphic to  $\mathcal{O}_K/\mathfrak{a}$ ; thus we obtain a natural map  $\pi_{\mathfrak{a}} : \Lambda^{r+1} \rightarrow (\mathcal{O}_K/\mathfrak{a})^{r+1}$ . For the lower bound sieve, when we have  $\mathfrak{a} = (p)$ , we denote this map by  $\pi_p$ . If  $\mathfrak{a} = \mathfrak{p}_1 \mathfrak{p}_2 \cdots \mathfrak{p}_k$  is square-free, then by the Chinese remainder theorem, we get an isomorphism

$$\varphi_{\mathfrak{a}} : (\mathcal{O}_K/\mathfrak{p}_1)^{r+1} \oplus (\mathcal{O}_K/\mathfrak{p}_2)^{r+1} \oplus \cdots \oplus (\mathcal{O}_K/\mathfrak{p}_k)^{r+1} \rightarrow (\mathcal{O}_K/\mathfrak{a})^{r+1},$$

such that the map  $\pi_{\mathfrak{a}}$  is compatible with the maps  $\pi_{\mathfrak{p}_i}$  for all  $i$ .

A *siftable set* is a triple  $(X, \mu, F)$ , consisting of a finite measure space  $(X, \mu)$ , and a map  $F : X \rightarrow Y$ , such that the composite map  $\pi_\alpha \circ F$  is measurable. This introduces a lot of flexibility, because  $X$  does not have to be a subset of  $\Lambda^{r+1}$ . The discussion at the beginning of this section shows how to construct a map  $F : X \rightarrow \Lambda^{r+1}$ , in the case where  $X$  is a finite subset of  $\mathbb{P}_K(K)$ . If  $\mu$  is the counting measure on  $X$ , then it is clear that the composite maps  $\pi_{\mathfrak{p}} \circ F$  are measurable, so this gives us a siftable set. Now if we have a finite morphism  $\varphi : U_K \rightarrow \mathbb{P}_K^r$ , we get a height  $H_\varphi$  on  $U_K$  in the usual way, and then we can look at a subset  $X$  of

$$B_K(x) = \{u \in U_K(K) : H_\varphi(u) \leq x\}.$$

Since the morphism  $\varphi$  is finite, it follows that the set  $X$  is finite, and so by composition with the map  $\mathbb{P}_K^r(K) \rightarrow \Lambda^{r+1}$  constructed above, we get a map  $F : X \rightarrow \Lambda^{r+1}$ . But we can go even further than this. If  $\varphi : V_K \rightarrow \mathbb{P}_K^r$  is a finite rational map, then by definition we have equivalence classes of pairs  $\langle U, \varphi_U \rangle$  such that  $\varphi_U : U \rightarrow \mathbb{P}_K^r$  is a finite morphism, and hence we can apply the argument above to this morphism. Given two pairs  $\langle U_1, \varphi_1 \rangle$  and  $\langle U_2, \varphi_2 \rangle$ , the compatibility condition gives us

$$H_{\varphi_1}|_{U_1 \cap U_2} = H_{\varphi_2}|_{U_1 \cap U_2},$$

where  $H_{\varphi_1}, H_{\varphi_2}$  are the heights corresponding to  $\varphi_1, \varphi_2$ . This allows us to safely speak of a height  $H_\varphi$  corresponding to the rational function  $\varphi$ , although it may

not be defined on all of  $V_K$ . However, so long as  $V_K$  is noetherian, the domain of the rational map has a finite cover by open sets  $U_1, \dots, U_N$ , and this turns out to be sufficient. In general, suppose  $(X, \mu)$  is a finite measure space, but we don't have a map  $F : X \rightarrow Y$ . If we have a finite cover  $X$  by the sets  $X_1, \dots, X_N$ , such that  $(X_i, \mu, F_i)$  is a siftable set, then we can use sub-additivity to extend the large sieve so that it applies to  $(X, \mu)$ , and thus we would like to call this a siftable set also. In particular, we can regard any subset  $X$  of

$$B_K(x) = \{t \in V_K(K) : t \in U \text{ for some } \langle U, \varphi_U \rangle, \text{ and } H_\varphi(t) \leq x\}$$

as a siftable set.

A *prime sieve support*  $\mathcal{L}^*$ , is a finite subset of the indexing set  $\mathcal{A}$ , and a *sieve support*  $\mathcal{L}$  is a subset of the power set of  $\mathcal{L}^*$ . Let  $\Sigma_K(Q)$  and  $\Sigma_K^1(Q; d, l)$  be defined as follows:

$$\Sigma_K(Q) = \{\mathfrak{p} \in \Sigma_K : N(\mathfrak{p}) < Q\}$$

$$\Sigma_K^1(Q; d, l) = \{\mathfrak{p} \in \Sigma_K(Q) : N(\mathfrak{p}) = q \text{ is prime, and } q \equiv d \pmod{l}\}.$$

In our general statement of the large sieve, we can take  $\mathcal{L}^*$  to be any subset of  $\Sigma_K(Q)$ , but in our application of we will use an appropriate subset of  $\Sigma_K^1(Q; d, l)$ . Note that any subset of  $\mathcal{L}^*$  can be viewed as a square-free product of these prime ideals. In particular we will define  $\mathcal{L}$ , to be the set of all square-free ideals  $\mathfrak{a}$ , such that  $N(\mathfrak{a}) < Q$  and for all  $\mathfrak{p}|\mathfrak{a}$  we have  $\mathfrak{p} \in \mathcal{L}^*$ . We will need compatibility, between the primes used in the lower bound sieve and the prime ideals used in the large sieve. If  $\mathfrak{p} \in \Sigma_K^1(Q; d, l)$  and  $N(\mathfrak{p}) = p$ , then  $p < Q$  hence these are the primes that will be used in the lower bound sieve. Now for each  $\alpha \in \mathcal{L}^*$  we choose a *sieving set*  $\Omega_\alpha \subset Y_\alpha$ , which may be completely arbitrary. If  $(X, \mu, F)$  is a siftable set under Kowalski's original definition, then the *sifted set* is

$$S(X, (\Omega_\alpha); \mathcal{L}^*) = \{x \in X | \pi_\alpha(F(x)) \notin \Omega_\alpha \forall \alpha \in \mathcal{L}^*\},$$

or more generally if  $(X, \mu)$  is a siftable set under our extended definition, and  $(X_i, \mu, F_i)$  is a finite cover by siftable sets under the original definition, then the sifted set is

$$S(X, (\Omega_\alpha); \mathcal{L}^*) = \{x \in X | \exists i \text{ such that } x \in X_i \text{ and } \pi_\alpha(F_i(x)) \notin \Omega_\alpha \forall \alpha \in \mathcal{L}^*\}.$$

Finally, we take  $\nu_{\mathfrak{p}}$  to be the uniform probability measure on  $Y_{\mathfrak{p}}$ .

## 2.1 The large sieve

We will prove a version of the large sieve for projective varieties, with the help of an older version for torsion free  $\mathcal{O}_K$ -modules proven by Serre [?], which we restate below in notation compatible with that described above (see also Zywin's paper [?]).

**Theorem 1.** *Let  $\Lambda$  be a torsion free  $\mathcal{O}_K$ -module with rank  $r + 1$  over  $\mathcal{O}_K$ . Let  $\|\cdot\|$  be a norm of  $\Lambda_{\mathbb{R}} = \mathbb{R} \otimes \Lambda$ . Let  $x \geq 1$  and  $Q > 0$  be real numbers, and for each  $\mathfrak{p} \in \Sigma_K$ , let  $\omega_{\mathfrak{p}} \in [0, 1]$ . Suppose that  $E \subset \Lambda$  satisfies the conditions:*

1.  $E$  is contained in a ball of radius proportional to  $x$ .
2. For every  $\mathfrak{p}$  with  $N(\mathfrak{p}) \leq Q$ , we have the inequality

$$|\pi_{\mathfrak{p}}| \leq (1 - \omega_{\mathfrak{p}})|\Lambda/\mathfrak{p}\Lambda|.$$

Then we have

$$|E| \ll_{K, \Lambda, \|\cdot\|} \frac{\max\{x^{(r+1)d}, Q^{2(r+1)}\}}{L(Q)}$$

where the implied constant depends only on  $K$ ,  $\Lambda$ , and  $\|\cdot\|$ , and where

$$L(Q) = \sum_{\mathfrak{a}} \prod_{\mathfrak{p}|\mathfrak{a}} \frac{\omega_{\mathfrak{p}}}{1 - \omega_{\mathfrak{p}}},$$

the sum being over all square-free ideals  $\mathfrak{a}$  with norm  $\leq Q$ .

In particular the constant does not depend on  $x$ ,  $Q$  or the numbers  $\omega_{\mathfrak{p}}$ , so long as  $E$  satisfies the conditions in the statement of the theorem. This means that it depends only on the sieve setting, which in our application remains the same for all  $l$ . Also it is important to realize, that the set  $E$  is actually the sifted set. The statement of the large sieve for projective varieties is as follows.

**Theorem 2** (Large Sieve). *Suppose we have a finite rational map  $\varphi : V_K \rightarrow \mathbb{P}_K^r$  and that  $V_K$  is noetherian. Let  $(X, \mu)$  be a siftable set for the sieve setting  $(\Lambda^{r+1}, \Sigma_K, (\pi_{\mathfrak{p}}))$ , such that  $X$  is contained in*

$$B_K(x) = \{t \in V_K(K) : t \in U \text{ for some } \langle U, \varphi_U \rangle, \text{ and } H_{\varphi}(t) \leq x\}.$$

Let  $\mathcal{L}^*$  be an arbitrary subset of  $\Sigma_K(Q)$  and let  $(\Omega_{\mathfrak{p}})$  be an arbitrary family of sieving sets. Then

$$|S(X, (\Omega_{\mathfrak{p}}); \mathcal{L}^*)| \ll_{K, r, \varphi} \frac{\max\{x^{(r+1)[K:\mathbb{Q}]}, Q^{2(r+1)}\}}{L(Q)} \quad (1)$$

where the implied constant depends only on  $K, r, \varphi$ , and

$$L(Q) = \sum_{\mathfrak{a} \in \mathcal{L}} \prod_{\mathfrak{p}|\mathfrak{a}} \frac{\nu_{\mathfrak{p}}(\Omega_{\mathfrak{p}})}{\nu_{\mathfrak{p}}(Y_{\mathfrak{p}} - \Omega_{\mathfrak{p}})} \quad (2)$$

where  $\mathcal{L}$  is the set of all square-free ideals  $\mathfrak{a}$ , such that  $N(\mathfrak{a}) \leq Q$  and for all  $\mathfrak{p}|\mathfrak{a}$  we have  $\mathfrak{p} \in \mathcal{L}^*$ .

*Remark.* The theorem is stated for relative heights. It is a simple matter to obtain a version of this theorem for absolute heights by observing that we have  $H = H_K^{1/[K:\mathbb{Q}]}$ , where  $H_K$  is the height relative to  $K$  and  $H$  is the absolute height. This means that the theorem remains correct for absolute heights if we replace the numerator of the fraction on the right hand side of (1) by  $\max\{x^{r+1}, Q^{2(r+1)}\}$ . It should also be noted that the implied constant depends only on data from the sieve setting,  $K, r$ , and on data from the siftable set,  $\varphi$ . In particular it does not depend in any way on  $\mathcal{L}^*$ .

*Proof.* By the noetherian condition we may assume that the domain of  $\varphi$  has a finite cover  $U_1, U_2, \dots, U_N$ . Let  $\varphi_i$  be the restriction of  $\varphi$  to  $U_i$ , and define

$$X_i = X \cap U_i(K) \quad \text{and} \quad F_i : X_i \rightarrow \Lambda^{r+1},$$

where  $\Lambda$  is the Minkowski lattice corresponding to  $K$ , and the map  $F_i$  is constructed as above. As noted above, Serre's theorem is stated for the sifted set, hence for convenience we define

$$E_i = S(X_i, (\Omega_{\mathfrak{p}}); \mathcal{L}^*).$$

Now  $\Lambda^{r+1}$  is a torsion free  $\mathcal{O}_K$ -module of rank  $r+1$ , but we must still verify that conditions 1 and 2 hold. The discussion at the beginning of the section shows that  $E_i$  is contained in a ball of radius proportional to  $x$ , which shows that condition 1 holds. As for condition 2, suppose that  $\mathfrak{p} \in \Sigma_K$  has norm  $N(\mathfrak{p}) \leq Q$ . If  $\mathfrak{p} \in \mathcal{L}^*$ , then

$$\pi_{\mathfrak{p}}(E_i) \subset Y_{\mathfrak{p}} - \Omega_{\mathfrak{p}}$$

where  $Y_{\mathfrak{p}} = (\Lambda/\mathfrak{p}\Lambda)^{r+1}$ . Even if  $\mathfrak{p} \notin \mathcal{L}^*$  we can still consider the images of  $E_i$ , which will be trivially contained in  $Y_{\mathfrak{p}}$ . Therefore, if we define

$$\omega_{\mathfrak{p}} = \begin{cases} \nu_{\mathfrak{p}}(\Omega_{\mathfrak{p}}) & \text{if } \mathfrak{p} \in \mathcal{L}^* \\ 0 & \text{otherwise} \end{cases}$$

then

$$\nu_{\mathfrak{p}}(\pi_{\mathfrak{p}}(E_i)) \leq 1 - \omega_{\mathfrak{p}}$$

in all cases. Since  $\nu_{\mathfrak{p}}$  is the uniform measure on  $Y_{\mathfrak{p}}$ , this gives us condition 2. By his theorem, we then have

$$|E_i| \ll_{\Lambda^{r+1}, \|\cdot\|} \frac{\max\{x^{(r+1)[K:\mathbb{Q}]}, Q^{2(r+1)}\}}{L(Q)}$$

where the implied constant depends only on  $\Lambda^{r+1}$  and  $\|\cdot\|$ , and

$$L(Q) = \sum_{\mathfrak{a}} \prod_{\mathfrak{p}|\mathfrak{a}} \frac{\omega_{\mathfrak{p}}}{1 - \omega_{\mathfrak{p}}}$$

where the sum is over all square-free ideals such that  $N(\mathfrak{a}) \leq Q$ . By the definition of  $\omega_{\mathfrak{p}}$  it follows that this definition of  $L(Q)$  is equivalent with the one in the statement of the theorem. In our case, the chosen norm is determined completely by  $\Lambda^{r+1}$ , and  $\Lambda^{r+1}$  depends only on  $K$  and on  $r$ , so the implied constant really only depends on  $K$  and  $r$ , and we indicate this by changing the subscript in the inequality. We get from  $E_i$  back to  $S(X_i, (\Omega_{\mathfrak{p}}); \mathcal{L}^*)$ , via the maps

$$V_K(K) \xrightarrow{\varphi} \mathbb{P}_K^r(K) \longrightarrow \Lambda^{r+1}.$$

The second map is injective, so it follows that  $|E_i| = |S(\varphi(X_i), (\Omega_{\mathfrak{p}}); \mathcal{L}^*)|$ , but the map  $\varphi$  does not have to be injective. Since  $\varphi$  is finite, at the very least we have

$$|S(X_i, (\Omega_{\mathfrak{p}}); \mathcal{L}^*)| \ll_{\varphi} |S(\varphi(X_i), (\Omega_{\mathfrak{p}}); \mathcal{L}^*)|,$$

and finally to get back to  $S(X, (\Omega_{\mathfrak{p}}); \mathcal{L}^*)$ , we use finite sub-additivity.  $\square$

## 2.2 The lower bound sieve

Iwaniec and Kowalski have proven the following small sieve result.

**Theorem 3.** *Let  $\kappa > 0$  and  $D > 1$ . There exist upper and lower-bound sieve coefficients  $(\lambda_d^{\pm})$  depending only on  $\kappa$  and  $D$ , supported on square-free integers  $< D$ , bounded by one in absolute value, with the following properties: for all  $s \geq 9\kappa + 1$  and  $Q^{9\kappa+1} < D$ , we have*

$$\begin{aligned} \int_{S(X, (\Omega_p); Q)} \alpha(u) d\mu(u) &< (1 + e^{9\kappa+1-s}) \prod_{p < Q} (1 - \nu_p(\Omega_p)) H + R^+(X; Q^s), \\ \int_{S(X, (\Omega_p); Q)} \alpha(u) d\mu(u) &> (1 - e^{9\kappa+1-s}) \prod_{p < Q} (1 - \nu_p(\Omega_p)) H - R^-(X; Q^s) \end{aligned}$$

provided that

$$\prod_{w \leq p < Q} \frac{1}{1 - \nu_p(\Omega_p)} \ll \left( \frac{\log Q}{\log w} \right)^{\kappa},$$

for all  $w$  and  $Q$  satisfying  $2 \leq w < Q < D$

In this theorem  $H$  and  $R^{\pm}(X; Q^s)$  are defined as follows

$$\begin{aligned} H &= \int_X \alpha(x) d\mu(x) & S_d(X; \alpha) &= \nu_d(\Omega_d) H + r_d(X; \alpha) \\ P(Q) &= \prod_{\substack{p \leq Q \\ l \in \mathcal{L}^*}} l & R^{\pm}(X; Q^s) &= \sum_{\substack{d \leq Q^s \\ d|P(Q)}} |\lambda_d^{\pm} r_d(X; \alpha)|. \end{aligned}$$

Thus  $H$  can be regarded as the main term and  $r_d$  or  $R^{\pm}$  are regarded as remainder terms. We apply this to the sieve setting  $(\Lambda^{r+1}, \Sigma_{\mathbb{Q}}, (\pi_p))$ , and siftable set  $(B_K(x), \mu, F)$ , where

$$B_K(x) = \{u \in \mathbb{P}_K^{r_K}(K) : H(u) \leq x\},$$

where  $\mu$  is the counting measure.

We also define

$$B_{\Lambda}(x) = \left\{ a \in \Lambda^{r+1} : \prod_{v \in S_{\infty}} \sup_i \|a_{iv}\|_v \leq x \right\},$$

where  $i$  ranges from 0 to  $r$ , and we denote the uniform probability measure on  $B_\Lambda(x)$  by  $P$ . By construction, the image of  $B_K(x)$  in  $\Lambda^{r+1}$  under the map  $F$  will be contained in  $B_\Lambda(x)$ , which is easier to work with.

The sifting sets are constructed by first choosing subsets  $\Omega_{\mathfrak{p}} \subset (\mathcal{O}_K/\mathfrak{p})^{r+1}$ , and looking at

$$A_p = \varphi_p\left(\prod_{\mathfrak{p}|p} A_{\mathfrak{p}}\right) \subset (\mathcal{O}_K/p)^{r+1}$$

where  $A_{\mathfrak{p}}$  is the complement of  $\Omega_{\mathfrak{p}}$ . Here the map  $\varphi_p$  comes from the Chinese remainder theorem, as discussed at the beginning of the section. Then we define  $\Omega_p$  to be the complement of  $A_p$ . If  $\nu_{\mathfrak{a}}$  denotes the uniform probability measure on  $(\mathcal{O}_K/\mathfrak{a})^{r+1}$  then the Chinese remainder theorem then gives us

$$1 - \nu_p(\Omega_p) = \prod_{\mathfrak{p}|p} (1 - \nu_{\mathfrak{p}}(\Omega_{\mathfrak{p}})). \quad (3)$$

In order to estimate the remainder terms  $R^\pm(X; Q^s)$ , we will need the following lemma.

**Lemma 1.** *Let  $\mathfrak{a}$  be an ideal, with  $N(\mathfrak{a}) \leq D^d$ . If*

$$S_{\mathfrak{a}} \subset (\mathcal{O}_K/\mathfrak{a})^{r+1} \quad \text{and} \quad S = B_\Lambda(x) \cap \pi_{\mathfrak{a}}^{-1}(S_{\mathfrak{a}})$$

*then*

$$\left| P(S) - \frac{|S_{\mathfrak{a}}|}{N(\mathfrak{a})^{r+1}} \right| \ll_{K,r} \begin{cases} \frac{D^2 \log x}{x} & \text{if } d = r = 1, \\ \frac{D^{d(r+1)}}{x} & \text{otherwise,} \end{cases} \quad (4)$$

*for sufficiently large  $x$ , where the implied constant depends only on  $r$ .*

*Proof.* Schanuel [?] has proven that if  $\mathfrak{a}$  is an ideal of  $\mathcal{O}_K$ , then the number of lattice points in  $B_\Lambda(x) \cap (\mathfrak{a}\Lambda)^{r+1}$  is

$$\varkappa \frac{x^{d(r+1)}}{N(\mathfrak{a})^{r+1}} + \begin{cases} O(x \log x) & \text{if } d = r = 1, \\ O(x^{d(r+1)-1}) & \text{otherwise,} \end{cases}$$

where  $\varkappa$  depends only on  $K$  and on  $m$ , and the same goes for the implied constant in the error term (see also Serre [?]). However, because of the method of proof, we can also use this estimate for the number of full copies of a system of representatives for  $(\Lambda/\mathfrak{a}\Lambda)^{r+1}$  that can be found in  $B_\Lambda(x)$ . If  $S \subset (\Lambda/\mathfrak{a}\Lambda)^{r+1}$ , then this gives us the estimate

$$\frac{|S|}{|S_{\mathfrak{a}}|} = \frac{|B_\Lambda(x)|}{N(\mathfrak{a})^{r+1}} + \begin{cases} O\left(\frac{|B_\Lambda(x)| \log x}{x}\right) & \text{if } d = r = 1, \\ O\left(\frac{|B_\Lambda(x)|}{x}\right) & \text{otherwise.} \end{cases}$$

If we multiply by  $\frac{|S_{\mathfrak{a}}|}{|B_\Lambda(x)|}$  and use the bound  $|S_{\mathfrak{a}}| \leq D^{d(r+1)}$ , then we obtain the result in the lemma.  $\square$



**Theorem 4.** *Given a homogeneous polynomial  $f$  in  $K[t_0, \dots, t_r]$ , let  $S$  be a finite subset of  $\Sigma_K^1$  satisfying the following properties*

1. *if  $f \equiv 0 \pmod{\mathfrak{p}}$ , then  $\mathfrak{p} \in S$ ,*
2. *if  $\mathfrak{p} \in S$  and  $\mathfrak{p}|p$ , where  $p$  is a prime in  $\mathbb{Z}$ , then  $\mathfrak{q} \in S$  for all  $\mathfrak{q}|p$ ,*
3. *if  $p$  ramifies in  $K$  and  $\mathfrak{p}|p$  then  $\mathfrak{p} \in S$ ,*

and for  $\mathfrak{p} \notin S$  let

$$\Omega_{\mathfrak{p}} = \{(t_0, \dots, t_r) \in \mathbb{F}_{\mathfrak{p}}^{r+1} : f(t_0, \dots, t_r) \equiv 0 \pmod{\mathfrak{p}}\}. \quad (5)$$

Let  $\mathcal{L}^* \subset \Sigma_K^1(Q) \setminus S$ . Then there exists  $\kappa$  such that

$$|\{u \in B_K(x) : \pi_{\mathfrak{p}}(F(u)) \notin \Omega_{\mathfrak{p}} \ \forall \mathfrak{p} \in \mathcal{L}^*\}| \gg_{K,r} \frac{x^{r+1}}{(\log Q)^{\kappa}},$$

so long as  $Q^{s(d(r+1)+1)} \leq x^{\frac{1}{2}}$  and  $s > 9\kappa + 1$ .

*Proof.* Let  $W_{\mathfrak{p}}$  denote the variety defined by  $f = 0$  over  $\mathbb{F}_{\mathfrak{p}}$ , and let  $q = |\mathbb{F}_{\mathfrak{p}}|$ . Trivially, there exists a constant  $\kappa_1$ , depending only on the degree of  $f$ , such that

$$|W_{\mathfrak{p}}(\mathbb{F}_{\mathfrak{p}})| \leq \kappa_1 q^{r-1}.$$

Since each point in  $\mathbb{P}^r(\mathbb{F}_{\mathfrak{p}})$  corresponds with  $q - 1$  points in  $\mathbb{F}_{\mathfrak{p}}^{r+1}$ , we obtain

$$\nu_{\mathfrak{p}}(\Omega_{\mathfrak{p}}) \leq \frac{\kappa_1}{N(\mathfrak{p})}.$$

Suppose that  $\mathfrak{p} \in S$  and lies over  $p$ . Then by (3) we have

$$1 - \nu_p(\Omega_p) = \prod_{\mathfrak{p}|p} \left(1 - \frac{1}{|\mathbb{F}_{\mathfrak{p}}|}\right) \sim \left(1 - \frac{\kappa_1}{p}\right)^d,$$

and it can be shown that

$$\prod_{w \leq p < Q} \left(1 - \frac{\kappa_1}{p}\right)^{-d} \ll \left(\frac{\log Q}{\log w}\right)^{d\kappa_1}$$

where the implied constant depends on the error term in the prime number theorem. Since  $|B_K(x)| \asymp |B_{\Lambda}(x)|$  then by applying lemma 1 to  $\mathfrak{a} = d\mathcal{O}_K$  and  $S = \Omega_d$ , we obtain

$$|r_d(B_K(x); \alpha)| \ll_{K,r} \begin{cases} x \log x D^2 & \text{if } d = r = 1, \\ x^r D^{d(r+1)} & \text{otherwise.} \end{cases}$$

For  $D = Q^s$ , this gives us

$$R^-(B_K(x); \alpha) \ll_{K,r} \begin{cases} x \log x Q^{3s} & \text{if } d = r = 1, \\ x^r Q^{s(d(r+1)+1)} & \text{otherwise,} \end{cases}$$

which will be an acceptable error if  $Q^{s(d(r+1)+1)} \leq x^{\frac{1}{2}}$ . The prime number theorem gives us

$$\prod_{p < Q} \left(1 - \frac{\kappa_1}{p}\right) \ll \frac{1}{(\log Q)^{d\kappa_1}},$$

and so by applying theorem 3 with  $\alpha(x) = 1$ , we obtain the result stated in the theorem with  $\kappa = d\kappa_1$ .  $\square$

### 3 General machinery

All definitions made in the previous section will be maintained, and we will supplement them with the following. For any group  $G$ , the set of conjugacy classes will be denoted by  $G^\#$ . If  $V$  is a scheme over a field  $k$ , then  $\overline{\eta}$  will denote a geometric point, and  $\overline{V}$  will denote the extension of  $V$  to a separable closure  $k^{\text{sep}}$ .

Now in particular if  $f : U \rightarrow V$  is a finite étale covering over  $k$  with arithmetic monodromy group  $G$  and geometric monodromy group  $G^g$ , then we get the following commutative diagram:

$$\begin{array}{ccccccc} 1 & \longrightarrow & \pi_1(\overline{V}, \overline{\eta}) & \longrightarrow & \pi_1(V, \overline{\eta}) & \longrightarrow & \text{Gal}(k^{\text{sep}}/k) \longrightarrow 1 \\ & & \downarrow & & \downarrow \rho & & \downarrow \\ 1 & \longrightarrow & G^g & \longrightarrow & G & \longrightarrow & G/G^g \longrightarrow 1 \end{array}$$

The important point is that we must consider this diagram both where  $k$  is a number field  $K$ , and where  $k$  as a finite field  $\mathbb{F}_{\mathfrak{p}}$ , which we obtain by reducing mod  $\mathfrak{p}$  for some  $\mathfrak{p} \in \Sigma_K$ . Moreover, we will have a family of coverings  $f_l : V_l \rightarrow V$  over  $K$ , parametrized by  $l$ , where  $l$  is a prime. In this case, we will denote the corresponding monodromy groups by  $G_l$  and  $G_l^g$  respectively, and the map  $\rho$  will be renamed  $\rho_l$ . For each  $l$ , when we reduce mod  $\mathfrak{p}$  to obtain the analogous situation over  $\mathbb{F}_{\mathfrak{p}}$ , we will also introduce  $\mathfrak{p}$  as a subscript:  $G_{l,\mathfrak{p}}$ ,  $G_{l,\mathfrak{p}}^g$  and  $\rho_{l,\mathfrak{p}}$ . The question arises, with  $l$  fixed, when do we have  $G_l \cong G_{l,\mathfrak{p}}$  or similarly  $G_l^g \cong G_{l,\mathfrak{p}}^g$ . This may not happen for all  $\mathfrak{p}$ , but in our application we manage to get sufficient control over this.

#### 3.1 The Chebotarev density theorem

The following theorem is based on a version of the Chebotarev density theorem originally due to Lang [?], which we sharpen with the help of a recent result by Kowalski [?]. Since it applies to more general situations than the one here, we will state it for a single Galois covering  $f : U \rightarrow V$ , and thus we will omit the subscripts  $l, \mathfrak{p}$ .

**Theorem 5.** *Let  $U, V$  be a smooth geometrically irreducible affine schemes over  $\mathbb{F}_q$  of dimension  $r \geq 1$ , and let  $f : U \rightarrow V$  be a finite étale covering, with generic*

Galois group  $G$  such that  $(|G|, q) = 1$ . If  $C \in G^\#$  is arbitrary, and we define

$$\Omega_C = \{t \in V(\mathbb{F}_q) : \rho(\text{Frob}_t) \in C\},$$

then

$$\left| \frac{|\Omega_C|}{q^r} - \frac{|C|}{|G|} \right| \ll |G|^{\frac{1}{2}} |G^\#|^{\frac{1}{2}} |C| q^{-\frac{1}{2}} \quad (6)$$

where the implied constant depends only on  $\bar{V}$ , and in particular it does not depend on  $q$  or on  $G$ .

*Proof.* Let  $g \in C$  be arbitrary. By orthogonality of characters we get

$$\begin{aligned} |\Omega_C| \frac{|G|}{|C|} &= \sum_{t \in V(\mathbb{F}_q)} \sum_{\chi} \chi(\rho(\text{Frob}_t)) \overline{\chi(g)} \\ &= |V(\mathbb{F}_q)| + \sum_{\chi \neq 1} \overline{\chi(g)} \sum_{t \in V(\mathbb{F}_q)} \chi(\rho(\text{Frob}_t)), \end{aligned}$$

and hence we have the inequality

$$\left| \frac{|\Omega_C|}{q^r} - \frac{|C|}{|G|} \right| \leq \frac{|C|}{|G|} \left| \frac{|V(\mathbb{F}_q)|}{q^r} - 1 \right| + \frac{|C|}{|G|} \frac{1}{q^r} \sum_{\chi \neq 1} \left| \sum_{t \in V(\mathbb{F}_q)} \chi(\rho(\text{Frob}_t)) \right|.$$

The result of Lang and Weil in [?], gives us a bound for the first term on the right of the form  $Aq^{-\frac{1}{2}}$ , where the constant  $A$  clearly has no dependence on  $G$ . As for the summation, if we apply proposition 5.1 (i) in [?] with  $l = l', \pi' = 1$ , and the irreducible representation  $\pi$ , chosen so that  $\chi = \text{Tr}(\pi)$ , then we obtain

$$\sum_{\chi \neq 1} \left| \sum_{t \in V(\mathbb{F}_q)} \chi(\rho(\text{Frob}_t)) \right| \leq C(\bar{V}) q^{r-\frac{1}{2}} |G| \sum_{\pi \neq 1} \dim \pi,$$

where  $C(\bar{V})$  is a constant depending only on  $\bar{V}$ . Then by using Cauchy's inequality, we get the result in the statement of the theorem.  $\square$

### 3.2 The general theorem

Let  $V$  be a geometrically irreducible affine scheme over  $K$  of dimension  $r \geq 1$ , birationally equivalent to  $\mathbb{P}_K^r$  via the rational map  $\varphi : V \rightarrow \mathbb{P}_K^r$ . Let  $\langle U, \varphi \rangle$  be a pair defining this rational map, let  $H_\varphi$  denote the corresponding absolute height on  $U$ , and define

$$B_K(x) = \{t \in V(K) : \exists \langle U, \varphi \rangle \text{ such that } t \in U(K) \text{ and } H_\varphi(t) \leq x\} \quad (7)$$

For each prime  $l$ , let  $V_l$  be a variety of dimension  $r$  over  $K$ , such that  $f_l : V_l \rightarrow V$  is a finite étale covering with generic Galois group  $G_l$ , and suppose that these groups satisfy the following property:

**Property 1.** *There exist constants  $\alpha, \beta$ , such that*

$$|G_l| \ll l^\alpha \text{ and } |G_l^\#| \ll l^\beta$$

*hold for all  $l$ , where the implied constants are absolute.*

Obviously the condition  $|G_l^\#| \ll l^\beta$  is not necessary in property 1, however, in our application it is possible to take  $\beta$  smaller than  $\alpha$ , giving us a sharper bound.

If  $t \in V(K)$ , then  $G_K = \text{Gal}(\overline{K}/K)$  acts on the fiber  $V_{l,t}$ , which gives us a homomorphism  $\rho_{l,t} : G_K \rightarrow G_{l,t}$ . To determine the  $t$  for which we have surjectivity on  $G_l$ , we use the following lemma of Jordan (see [?]):

**Lemma 2** (Jordan). *Given a finite group  $G$  and a subgroup  $H \subset G$ , if  $H \cap C \neq \emptyset$  for every  $C \in G^\#$ , then  $H = G$ .*

By this lemma, it suffices to show that the intersection with all conjugacy classes in  $G_l$  are non-empty. This is accomplished by applying the large sieve 2 to each conjugacy class in  $G_l$ , with  $\mathcal{L}^*$  taken to be an appropriate subset of  $\Sigma_K^1(Q; d, l)$ , where  $d$  is the common determinant of that conjugacy class. The constant in the large sieve does not depend on  $l$  because the sieve setting and  $\varphi$  remain the same for all  $l$ . To get a useful estimate though, we will then need the Chebotarev density theorem 5; thus we need  $V_{\mathfrak{p}}$  to be geometrically irreducible and  $G_l \cong G_{l,\mathfrak{p}}$  for all  $\mathfrak{p} \in \mathcal{L}^*$ . Since  $l$  is fixed, theorem 9.7.7 of [?] shows that there are only a finite number for  $\mathfrak{p}$  for which  $V_{\mathfrak{p}}$  is not geometrically irreducible. In our application, we apply this together with known results for the moduli spaces of abelian varieties, in order to get the following property to hold.

**Property 2.** *There exists a constant  $\delta$ , such that*

$$|\{\mathfrak{p} \in \Sigma_K : G_{l,\mathfrak{p}} \not\cong G_l\}| \ll l^\delta$$

*holds for all  $l$ , where the implied constant is absolute.*

With  $\mathcal{L}^*$  chosen, if we take  $\mathfrak{p} \in \mathcal{L}^*$  and reduce  $V_l \rightarrow V \bmod \mathfrak{p}$ , we may no longer have a finite étale morphism. To deal with this, we define

$$\Omega(\mathcal{L}^*) = \{t \in V(K) : V_{l,\mathfrak{p},t} \rightarrow V_{\mathfrak{p},t} \text{ is finite étale for all } \mathfrak{p} \in \mathcal{L}^*\}.$$

In this way by removing the fibers over  $\Omega(\mathcal{L}^*)$  before reducing  $V_l \rightarrow V \bmod \mathfrak{p}$ , we obtain a finite étale morphism  $V_{l,\mathfrak{p}} \rightarrow V_{\mathfrak{p}}$  defined over  $\mathbb{F}_{\mathfrak{p}}$ . We also define a corresponding siftable set

$$C_K(x) \subset \{t \in B_K(x) : t \notin \Omega(\mathcal{L}^*)\}, \quad (8)$$

which we want to have the following property:

**Property 3.** *There exists a constant  $\kappa$  so that  $|C_K(x)| \gg_{K,r} \frac{x^{r+1}}{(\log x)^\kappa}$ .*

*Remark.* It is important to note that the definition of  $C_K(x)$  depends on the choice of  $\mathcal{L}^*$ , and thus on  $Q$ . So, in order for this property to make sense, there must be a fixed relationship between  $x$  and  $Q$ , specifically we will take  $Q$  to be an appropriate power of  $x$ . In the case where equality holds in (8), Theorem 4 shows that if the image of  $C_K(x)$  in  $\mathbb{P}_K^r(K)$  can be covered by a constructible set not containing the generic point, then this property will hold so long as  $Q^{s(d(r+1)+1)} \leq x^{\frac{1}{2}}$  and  $s > 9\kappa + 1$ . Note that in order for theorem 4 to be applied here we need  $V$  to be unirational, and a rational map  $\varphi : V \rightarrow \mathbb{P}_K^r$  is required to define a height on  $V$ , hence we will need to assume that  $V$  is a rational variety in order to use both of these.

The Chebotarev density theorem also requires  $(|G_l|, p) = 1$ , and this is satisfied if  $p > |G_l|$ . When this condition is included in the definition of  $\mathcal{L}^*$ , it follows that the primes in  $\mathcal{L}^*$  have a lower bound for their norm tending to infinity as  $x$  does. This is important, because it means that  $\limsup_x C_K(x)$  is the entire set of rational points.

In our application, we will need an effective version of Hilbert irreducibility for a single value of  $l$ , so we state it in the following lemma:

**Lemma 3** (Hilbert irreducibility). *Let  $K$  be a number field, and let  $V$  be a smooth geometrically irreducible affine scheme over  $K$ , birationally equivalent to  $\mathbb{P}_K^r$  for some  $r \geq 1$ . Fix a prime  $l$ , and let  $f : V_l \rightarrow V$  be a Galois covering with group  $G_l$ . Suppose 3 is satisfied with  $Q = x^\varepsilon$ , and define*

$$E_{K,l}(x) = \{t \in C_K(x) : \rho_{l,t}(G_K) \subsetneq G_l\}.$$

*Then for sufficiently large  $x$ , we have*

$$\frac{|E_{K,l}(x)|}{|C_K(x)|} \ll_{\varphi,K,r,l} |G_l| |G_l^\#| \cdot l \frac{(\log x)^{\kappa+1}}{x^\varepsilon}. \quad (9)$$

*Proof.* Since the absolute height is used in the definition of  $B_K(x)$  we use the adjustment to the large sieve made in the remark just below theorem 2. In our application of the large sieve we will fix  $l$  and  $C \in G_l^\#$ , and for convenience we will take  $Q = x^\varepsilon$ , for some  $\varepsilon \in (0, 1)$ . If theorem 4 is used to get property 3 to hold, then we will need  $\varepsilon \leq \frac{1}{2s(d(r+1)+1)}$ , where  $s > 9\kappa + 1$ . However, it is possible in special cases to take  $\varepsilon = \frac{1}{2}$ . Let  $d$  be the determinant of  $C$ , and define

$$\mathcal{L}^* = \{\mathfrak{p} \in \Sigma_K^1(Q; d, l) : N(\mathfrak{p}) > |G_l| \text{ and } G_{l,\mathfrak{p}} \cong G_l\}.$$

and for all  $\mathfrak{p} \in \mathcal{L}^*$  define

$$\Omega_{\mathfrak{p},C} = \{t \in V_{\mathbb{F}_\mathfrak{p}}(\mathbb{F}_\mathfrak{p}) : \rho_{l,\mathfrak{p}}(\text{Frob}_t) \in C\}.$$

We now want to estimate the right hand side of (2) in a useful way, with the help of the Chebotarev density theorem. Specifically, using (6), we get the following result

$$L(Q) \geq \sum_{\mathfrak{p} \in \mathcal{L}^*} \nu_{\mathfrak{p}}(\Omega_{\mathfrak{p},C}) \gg \sum_{\mathfrak{p} \in \mathcal{L}^*} \frac{|C|}{|G_l|}. \quad (10)$$

where the last inequality holds only for sufficiently large  $x$ , and the implied constant does not depend on  $G_l$  (in fact we can take the constant to be  $\frac{1}{2}$ ). We now must obtain an underestimate for  $|\mathcal{L}^*|$ . Since  $l$  is fixed, Grothendiek's theorem 9.7.7 of [?] shows that the set of  $\mathfrak{p} \in \Sigma_K$  for which  $G_{l,\mathfrak{p}} \not\cong G_l$  is finite, and hence the difference

$$|\Sigma_K^1(Q; d, l)| - |\mathcal{L}^*|$$

is bounded. This means that we can use Siegel-Walfisz theorem to get an estimate for  $|\mathcal{L}^*|$ . Specifically, by recalling that  $Q = x^\varepsilon$ , we have

$$|\mathcal{L}^*| \gg_{K,l} \frac{x^\varepsilon}{l \cdot \log x}$$

for sufficiently large  $x$ . Alternately, for an effective result, we can adjust the argument using the pigeon hole principle in a way completely analogous to Zywinia's version for elliptic curves [?]. In any case, by applying this to (10) we get the following underestimate for  $L(Q)$ ,

$$L(Q) \gg_{K,l} \frac{|C|}{|G_l|} \frac{x^\varepsilon}{l \cdot \log x},$$

and then using this estimate of  $L(Q)$  together with (1) gives us the following upper bound for the sifted set

$$|S(C_K(x), (\Omega_{\mathfrak{p},C}); \mathcal{L}^*)| \ll_{\varphi,K,r,l} \frac{|G_l|}{|C|} \cdot l \frac{\log x}{x^\varepsilon} x^{r+1} \quad (11)$$

In the future we will use the following shorthand for the sifted set, to make the remaining argument easier to follow

$$Y_C(x) = S(C_K(x), (\Omega_{\mathfrak{p},C}); \mathcal{L}^*).$$

By property 3 we have  $|C_K(x)| \gg \frac{x^{r+1}}{(\log x)^\kappa}$ , and so by (11)

$$\frac{|Y_C(x)|}{|C_K(x)|} \ll_{\varphi,K,r,l} \frac{|G_l|}{|C|} \cdot l \frac{(\log x)^{\kappa+1}}{x^\varepsilon} \quad (12)$$

for any conjugacy class  $C$  of  $G_l$ . Now by lemma 2  $E_{K,l}(x) \subset \bigcup_{C \in G_l^\#} Y_C(x)$ , so

$$\frac{|E_{K,l}(x)|}{|C_K(x)|} \leq \sum_{C \in G_l^\#} \frac{|Y_C(x)|}{|C_K(x)|}. \quad (13)$$

By combining these last two estimates, and using the trivial bound  $|C| \geq 1$  for each conjugacy class, we obtain the result in the lemma.  $\square$

*Remark.* Grothendiek's theorem 9.7.7 is proven in such a soft way, that it is not clear if it gives any control over the size of the finite set of bad primes as  $l$  varies. This is the only reason that the implied constant may depend on  $l$ . All other results used in the proof, show the dependence on  $l$  explicitly.

To prove a similar result that holds for all  $l$ , requires stronger methods. In particular if we define

$$E_K(x) = \{t \in C_K(x) : \rho_{l,t}(G_K) \subsetneq G_l \text{ for some } l\},$$

then we will also need the following lemma:

**Property 4.** *There exists a constant  $\gamma$  s.t.*

$$E_K(x) \subset \bigcup_{l \ll (\log x)^\gamma} E_{K,l}(x).$$

Now, here is the precise statement of our general result.

**Theorem 6.** *Let  $K$  be a number field, and let  $V$  be a smooth geometrically irreducible affine scheme over  $K$ , birationally equivalent to  $\mathbb{P}_K^r$  for some  $r \geq 1$ . For each prime  $l$ , let  $f : V_l \rightarrow V$  be a Galois covering with group  $G_l$ . Suppose that properties 2 and 4 are satisfied, that 3 is satisfied with  $Q = x^\varepsilon$ , and that  $\alpha, \beta$  are specific values for which property 1 is satisfied. Then for sufficiently large  $x$ , we have*

$$\frac{|E_K(x)|}{|C_K(x)|} \ll_{\varphi, K, r} \frac{(\log x)^{(\alpha+\beta+2)\gamma+\kappa+1}}{x^\varepsilon}.$$

*Proof.* The proof begins in the same way as it does for lemma 3, except that we now need properties 2 and 4 in order to get the estimate for  $|\mathcal{L}^*|$ . Using these properties together with the bound on  $|G_l|$  from property 1 we obtain the bound

$$|\Sigma_K^1(Q; d, l)| - |\mathcal{L}^*| \ll (\log x)^{\max\{\alpha, \delta\}\gamma},$$

where the implied constant is absolute. Comparing this with the estimate that Siegel-Walfisz gives us for  $|\Sigma_K^1(Q; d, l)|$  shows that this is indeed smaller for sufficiently large  $x$ . Continuing with the rest of the proof of lemma 3, we obtain (9) just as before, except with the constant not depending on  $l$ . We can now rewrite it purely in terms of  $x$  and  $l$  by using the bounds in property 1. Specifically, we get

$$\frac{|E_{K,l}(x)|}{|C_K(x)|} \ll_{\varphi, K, r} l^{\alpha+\beta+1} \frac{(\log x)^{\kappa+1}}{x^\varepsilon}. \quad (14)$$

Now by property 4 we have

$$\frac{|E_K(x)|}{|C_K(x)|} \leq \sum_{l \ll (\log x)^\gamma} \frac{|E_{K,l}(x)|}{|C_K(x)|},$$

and by applying (14) to this, it follows that

$$\frac{|E_K(x)|}{|C_K(x)|} \ll_{\varphi, K, r} \sum_{l \ll (\log x)^\gamma} l^{\alpha+\beta+1} \frac{(\log x)^{\kappa+1}}{x^\varepsilon} \ll_{\varphi, K, r} \frac{(\log x)^{(\alpha+\beta+2)\gamma+\kappa+1}}{x^\varepsilon}.$$

□

### 3.3 The method of Cojocaru and Hall

If we have a setup meeting the conditions of theorem 6, then the result gives us surjectivity of the maps

$$\rho_{l,t} : G_K \rightarrow G_l$$

for all  $l$ , and almost all  $t$  in the asymptotic sense that

$$\lim_{x \rightarrow \infty} \frac{|E_K(x)|}{|C_K(x)|} = 0.$$

This is true, whatever the groups  $G_l$  may happen to be. However, the determination of the groups  $G_l$  is not always an easy task. In this regard, we would like to consider the method that Cojocaru and Hall used for determining these groups in the case of elliptic curves [?], and show how this method generalizes. Suppose that we have auxiliary galois coverings  $X_l \rightarrow X$ , with known monodromy groups  $\Gamma_l$ , along with a commutative diagram

$$\begin{array}{ccc} V_l & \xrightarrow{\psi_l} & X_l \\ \downarrow & & \downarrow \\ V & \xrightarrow{\psi} & X \end{array}$$

where the horizontal maps are dominant covering maps. Our assumptions about  $V$  mean that  $X$  must be unirational. We will mainly be interested in the case where  $V_l$  is constructed by taking an irreducible component of  $V \times X_l$ . The commutative diagram gives us  $G_l \subset \Gamma_l$  at the very least, and we want to show that we have equality  $G_l = \Gamma_l$  for sufficiently large  $l$ , and possibly for all  $l$ . The trick is to look at an intermediate covering  $Y_l \rightarrow X$ , which is constructed as follows. Let  $K(\eta), K(\eta_l), K(\xi), K(\xi_l)$  be the function fields of  $V, V_l, X, X_l$ , so that

$$G_l = \text{Gal}(K(\eta_l)/K(\eta)) \quad \text{and} \quad \Gamma_l = \text{Gal}(K(\xi_l)/K(\xi)).$$

If we look at  $K(\xi_l)^{G_l}$  this defines an intermediate extension of function fields that corresponds to an intermediate covering  $X_l \rightarrow Y_l \rightarrow X$ . In the étale case, the existence of  $Y_l$  follows from proposition 3.1 in Exposé V of [?]. Moreover, we get a map  $V_l \rightarrow Y_l$  by composition, which then factors through the map  $V_l \rightarrow V$ , giving us the diagram:

$$\begin{array}{ccccc} V_l & & \xrightarrow{\psi_l} & & X_l \\ \downarrow & & & \swarrow & \downarrow \\ & & Y_l & & \\ \uparrow & & \searrow & & \uparrow \\ V & & \xrightarrow{\psi} & & X \end{array}$$



In the case of elliptic curves,  $X_l, X$  are the moduli schemes and we have

$$\Gamma_l = \mathrm{GL}_2(l)/\langle \pm 1 \rangle \quad \text{and} \quad \Gamma_l^g = \mathrm{SL}_2(l)/\langle \pm 1 \rangle$$

(see Katz [?]). Cojocaru and Hall take  $V$  to be a curve, and then claim that the genus of  $Y_l$  cannot be larger than the genus of  $V$ . This would be true if we knew that  $V \rightarrow Y_l$  was unramified, but it is not clear from the construction that this is the case.

It is worthwhile to consider the special case of Cojocaru and Hall, before considering the general version. The genus of  $X_l$  is known by the standard theory of  $l$ -level structure for elliptic curves, and the genus of  $Y_l$  must be determined. By construction, the covering map  $X_l \rightarrow Y_l$  has degree equal to  $|G_l|$ . Since the genus of  $X_l$  is roughly proportional to  $|\Gamma_l|$  for sufficiently large  $l$ , we expect intuitively that the genus of  $Y_l$  is roughly proportional to  $[\Gamma_l : G_l]$ . Certainly Table 2.1, which Cojocaru and Hall obtained by applying the Riemann Hurwitz formula to the covering  $X_l \rightarrow Y_l$ , vindicates this viewpoint. If it can be shown that  $[\Gamma_l : G_l]$  tends to infinity with  $l$ , then an upper bound on the genus of  $Y_l$  gives us  $G_l = \Gamma_l$  for sufficiently large  $l$ . Cojocaru and Hall do this in the case of the geometric monodromy groups  $G_l^g, \Gamma_l^g$ , by looking at the images in  $\mathrm{PSL}_2(\mathbb{F}_l)$  and using Serre's theorem concerning its structure.

In light of the previous consideration, however, it would seem that it is really  $[\Gamma_l : G_l]$  that is important and not the genus of  $Y_l$ . In fact  $[\Gamma_l : G_l]$  is the degree of the covering  $Y_l \rightarrow X$ , and since  $\psi : V \rightarrow X$  factors through this covering, it becomes clear that  $[\Gamma_l : G_l]$  must divide  $\deg \psi$ , and hence  $[\Gamma_l : G_l]$  remains bounded. If we can choose  $\psi : V \rightarrow X$  well enough so that  $\deg \psi = 1$ , then we have  $\Gamma_l = G_l$  for all  $l$ . This requires  $X$  to be a rational variety, which is certainly true for the moduli space of elliptic curves, and also for abelian surfaces (see [?]). This will be assumed in our application.

## 4 Application to abelian varieties

Let  $K$  be a number field, let  $V$  be a smooth geometrically irreducible affine scheme over  $K$  of dimension  $r \geq 1$ , birationally equivalent to  $\mathbb{P}_K^r$  via the rational map  $\varphi : V \rightarrow \mathbb{P}_K^r$ . Let  $K(V)$  be the function field of  $V$ , and let  $A$  be a principally polarized abelian variety over  $K(V)$  of dimension  $g$ . We obtain a map  $\psi : V \rightarrow X$ , where  $X$  is the Siegel moduli space of level 1, by sending  $t \in V$  to the point in  $X$  representing  $A_t$ , where  $A_t$  is the fiber above  $t$ . We also have a covering  $X_l \rightarrow X$  of degree  $|\mathrm{GSp}_{2g}(l)/\langle \pm 1 \rangle|$ . We have to mod out by  $\langle \pm 1 \rangle$ , as explained by Katz in [?] for the case  $g = 1$ . Then we define  $V_l$  to be an irreducible component of  $V \times_X X_l$ , so that we get the following commutative diagram.

$$\begin{array}{ccc} V_l & \longrightarrow & X_l \\ \downarrow & & \downarrow \\ V & \xrightarrow{\psi} & X \end{array}$$

As mentioned in the discussion of the method of Cojocaru and Hall, we will assume  $\psi : V \rightarrow X$  to be a dominant map with degree 1. This is not possible for all Siegel moduli spaces. It is known that the level 1 moduli spaces of principally polarized abelian varieties are unirational when  $g \leq 5$ , that they are of general type when  $g \geq 7$ , and the problem is still open in the case  $g = 6$  (see [?]). It is also known that when the level is  $\geq 4$ , moduli spaces of principally polarized abelian surfaces are of general type (see [?]). In the case of abelian surfaces it is known that the level 1 moduli space of polarized abelian surfaces is rational or unirational if the degree of polarization is 1, 2, 3, 4, 5, 7, or 9 (see [?]), in particular for the principally polarized case it is actually rational (see [?]). In the case of elliptic curves the analogous result is classical. Henceforth, we assume that we are in the cases  $g = 1$  or  $2$ , so that the rationality assumption holds.

Let  $G_K = \text{Gal}(\overline{K}/K)$ , fix  $t \in V(K)$  and consider the representation  $\rho_{l,t} : G_K \rightarrow \text{GL}(A_t[l])$ , where  $A_t[l]$  denotes the  $l$ -torsion of  $A_t$ . This representation induces a representation  $\rho_{l,t} : G_K \rightarrow G_l$ . However, whereas  $\text{GL}(A_t[l]) \cong \text{GSp}_{2g}(l)$ , we have  $G_l \cong \text{GSp}_{2g}(l)/\langle \pm 1 \rangle$  as noted above. As a consequence of this, it is easy to pass results from  $\text{GL}(A_t[l])$  to  $G_l$ , but to go backwards we will need the following lemma.

**Lemma 4.** *Given the exact sequence,*

$$1 \longrightarrow \langle \pm 1 \rangle \longrightarrow \text{GSp}_{2g}(l) \xrightarrow{\varphi} \text{GSp}_{2g}(l)/\langle \pm 1 \rangle \longrightarrow 1 \quad (15)$$

*let  $G$  be a subgroup of  $\text{GSp}_{2g}(l)$  s.t.  $\varphi|_G$  is surjective. Then  $G = \text{GSp}_{2g}(l)$ .*

*Proof.* It suffices to show that  $-1 \in G$ . If  $x \in \text{GSp}_{2g}(l)$ , then  $x$  or  $-x \in G$ . In particular

$$\begin{pmatrix} & -I_g \\ I_g & \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} & I_g \\ -I_g & \end{pmatrix}$$

is in  $G$ . But by squaring both of these matrices, it follows that  $-1 \in G$ .  $\square$

If  $H_\varphi$  is the absolute height on  $V$  corresponding to the map  $\varphi$ , then we define  $B_K(x)$  by equation (7). As for  $C_K(x)$ , we note that  $A_t$  must be non-singular and have good reduction at all primes  $\mathfrak{p} \in \mathcal{L}^*$ . By applying theorem 9.7.7 of [?] to  $A \rightarrow V$ , we obtain a constructible subset  $\Omega$  of  $V$  not containing the generic point, such that for any  $t \notin \Omega$  the fiber over  $t$  is non-singular. The image of this set under the rational map  $V \rightarrow \mathbf{P}_K^r$  is also constructible, and can be covered by varieties, which are defined by homogeneous polynomials in  $r + 1$  variables. Using theorem 4, we obtain a lower bound for the number of  $t \in B_K(x)$  for which  $A_t$  is non-singular and has good reduction mod  $\mathfrak{p}$  for all  $\mathfrak{p} \in \mathcal{L}^*$ . So far this enables us to take

$$C_K(x) = \left\{ t \in B_K(x) : \begin{array}{l} A_t \text{ is non-singular and has} \\ \text{good reduction mod } \mathfrak{p} \ \forall \mathfrak{p} \in \mathcal{L}^* \end{array} \right\}.$$

However, we also need to exclude values of  $t$  for which  $\text{End}(A_t) \neq \mathbb{Z}$ . In the elliptic curve case, this amounts to removing curves with complex multiplication. It would be nice to say that we simply need to remove the PEL varieties of lower dimension from the full Siegel moduli space, however, in general it is not clear that we have a finite list of possible endomorphism rings, once  $K$  has been fixed. So instead, we fix  $l$ , and apply the Hilbert irreducibility lemma 3 to obtain a subset of  $C_K(x)$ , for which  $\rho_{l,t}(G_K)$  is all of  $G_l$ . We then apply lemma 4 to lift to  $\text{GSp}_{2g}(\mathbb{F}_l)$ , and a result of Vasiu [?] to lift to  $\text{GSp}_{2g}(\mathbb{Z}_l)$ . This says that the number of  $t \in C_K(x)$  for which the galois representation on the Tate module associated to  $A_t$  is not surjective is bounded by a constant times

$$|C_K(x)| |G_l| |G_l^\#| \cdot l \frac{(\log x)^{\kappa+1}}{x^\varepsilon}.$$

Now Faltings [?] has shown that  $\text{End}(A_t) \otimes \mathbb{Z}_l \cong \text{End}_{G_K}(T_l A_t)$ , which means that if  $\text{End}(A_t) \neq \mathbb{Z}$ , then  $\text{End}_{G_K}(T_l A_t)$  contains a non-central element, and so the Galois representation on  $T_l A_t$  cannot be surjective. It follows that

$$|\{t \in C_K(x) : \text{End} A_t = \mathbb{Z}\}| \geq |C_K(x)| \left(1 - c |G_l| |G_l^\#| \cdot l \frac{(\log x)^{\kappa+1}}{x^\varepsilon}\right),$$

where  $c$  is the constant we get from lemma 3. Furthermore, if property 3 holds for  $C_K(x)$ , then it will hold for  $\{t \in C_K(x) : \text{End} A_t = \mathbb{Z}\}$  for sufficiently large  $x$ . To avoid complicating notation further, we simply redefine  $C_K(x)$  as follows

$$C_K(x) = \left\{ t \in B_K(x) : \begin{array}{l} \text{End}(A_t) = \mathbb{Z}, \text{ and } A_t \text{ is non-singular} \\ \text{and has good reduction mod } \mathfrak{p} \ \forall \mathfrak{p} \in \mathcal{L}^* \end{array} \right\}. \quad (16)$$

To show that property 4 holds using this data, we use the theorem of Masser and Wüstholz [?] in the case  $g = 1$ , and a generalization of it by Kawamura [?] in the case  $g = 2$ .

**Theorem 7** (Kawamura). *Let  $A$  be a principally polarized abelian surface over a number field of degree  $d$  with  $\text{End}_{\overline{K}}(A) \cong \mathbb{Z}$ . Let  $D(K)$  be the discriminant of  $K$ , and  $h(A)$  be the Faltings height of  $A$ . Then there exist constants  $c, \gamma$ , such that for any prime  $l$  satisfying*

$$l > \max\{D(K), c(\max\{3840d, h(A)\})^\gamma\},$$

*we have  $\rho_l(G_K) = \text{GSp}_4(l)$*

Kawamura tries to use the Main Theorem in Kleidman and Liebeck [?], however, that theorem applies only when the dimension is  $> 12$ . This condition is only needed for maximality, and so by not making any reference to maximality, Aschbacher's theorem can be applied instead (see Theorem 1.2.1 in [?]). There is another mistake in Kawamura's proof, specifically he claims that tables 3.5.A-H in [?] indicate that  $\mathcal{S}$  is empty, which is not true. This is also not a huge problem because the groups in  $\mathcal{S}$  can be dealt with in the same way that he deals with  $2^{1+4}.\text{O}_4^-(2)$ . With these minor changes, the proof of Theorem 7 is valid.

**Theorem 8.** *Property 4 holds for the groups  $G_l$  for  $C_K(x)$  as defined by equation (16) in the cases  $g = 1$  and  $2$ .*

*Proof.* Let  $t \in B_K(x)$ , be a point for which  $\rho_{l,t} : G_K \rightarrow G_l$  is not surjective, so that  $E_{K,l}$  is non-empty. Then the representation  $\rho_{l,t} : G_K \rightarrow \mathrm{GL}(A_t[l])$  is also not surjective, and so by Kawamura's theorem (which applies by the definition of  $C_K(x)$ ), there exist constants  $c, \gamma$  such that

$$l \leq \max\{D(K), c(\max\{3840d, h(A_t)\})^\gamma\} \ll h(A_t)^\gamma.$$

It is well known (see [?] for example) from the theory of heights that  $h(A_t) = h(t) + O(1)$ , where the height on the right is the absolute logarithmic height on  $\mathbb{P}_{\overline{\mathbb{Q}}}^r$ , i.e.  $h(t) = \log H(t)$ . The theory of height also gives us  $H_\varphi(t) = H(t) + O(1)$ , so it follows that

$$l \ll (\log H_\varphi(t))^\gamma \leq (\log x)^\gamma,$$

and therefore  $E(x)$  will be contained in the finite union  $\bigcup_{l \ll (\log x)^\gamma} E_{K,l}$ .  $\square$

We are now in a position to prove the following theorem

**Theorem 9.** *Let  $K$  be a number field, let  $V$  be a smooth geometrically irreducible affine scheme over  $K$  of dimension  $r = \binom{g+1}{2}$ , birationally equivalent to  $\mathbb{P}_K^r$  via the rational map  $\varphi : V \rightarrow \mathbb{P}_K^r$ . Let  $K(V)$  be the function field of  $V$ , and let  $A$  be a principally polarized abelian variety over  $K(V)$  of dimension  $g = 1$  or  $2$ . Let  $C_K(x)$  be defined by equation (16), and let*

$$E'_K(x) = \{t \in C_K(x) : \rho_{l,t}(G_K) \subsetneq \mathrm{GSp}_{2g}(l)\}.$$

*Then*

$$\lim_{x \rightarrow \infty} \frac{|E'_K(x)|}{|C_K(x)|} = 0.$$

*Proof.* Using the construction above, for each  $l$  we get a Galois covering  $V_l \rightarrow V$  with group  $G_l = \mathrm{GSp}_{2g}(l)/\langle \pm 1 \rangle$ . We have shown that property 4 holds in the cases  $g = 1$  and  $2$ . Under the assumption that the morphism  $\psi : V \rightarrow X$  is a degree 1 covering map, which exists since  $X$  is rational, we have  $G_l \cong \Gamma_l$  as explained in subsection 3.3. To obtain  $G_{l,\mathfrak{p}} \cong G_l$ , we apply reduction mod  $\mathfrak{p}$  to everything in the diagram and show that  $G_{l,\mathfrak{p}} \cong \Gamma_{l,\mathfrak{p}} \cong \Gamma_l$ . It is known that the level  $l$  Siegel moduli spaces have irreducible geometric fibers over  $\mathrm{Spec}(\mathbb{Z}[\zeta_l, \frac{1}{l}])$ , (see [?] for example). If  $\mathfrak{p}$  is a prime over one of the primes in  $\mathrm{Spec}(\mathbb{Z}[\zeta_l, \frac{1}{l}])$ , then we have  $\Gamma_{l,\mathfrak{p}} \cong \Gamma_l$ . We have already pointed out that  $V$  itself can be handled with theorem 9.7.7 of [?]. Taken together it follows that property 2 applies with  $\delta = 1$ .

For arbitrary dimension  $g$  the order formulas in [?] together with the upper bounds for  $|G_l^\#|$  in [?] (see also [?]), show that property 1 holds with  $\alpha = 2g^2 + g + 1$  and  $\beta = g + 1$ , so by applying our general Theorem 6 with these values, we obtain

$$\frac{|E_K(x)|}{|C_K(x)|} \ll_{F,K,r} \frac{(\log x)^{2(g^2+g+2)\gamma+\kappa+1}}{x^\varepsilon}. \quad (17)$$

In particular if  $g = 1$  or  $g = 2$  the factor in front of  $\gamma$  is 8 or 16 respectively. If theorem 4 is used to get property 3, then this estimate holds with

$$\varepsilon \leq \frac{1}{2s(d(r+1)+1)}$$

where  $s > 9\kappa + 1$  can be taken as a fixed number (independent of  $x$ ). However, as the example below shows, it may be possible to improve on the value of  $\varepsilon$ . If we view  $\rho_{l,t}$  as a representation on the  $l$ -torsion of  $A_t$  then  $\rho_{l,t}(G_K) \subset \mathrm{GSp}_4(l)$ . But consider the induced map  $\rho_{l,t} : G_K \rightarrow G_l$ , then  $\rho_{l,t}(G_K) \subset \mathrm{GSp}_4(l)/\langle \pm 1 \rangle$ . The two images are related by the exact sequence (15), hence  $t \in E'_K(x)$  implies that  $t \in E[x]$  by Lemma 4. It follows that we can replace  $E_K(x)$  with  $E'_K(x)$  in the estimates above, and so taking the limit  $x \rightarrow \infty$  proves the theorem.  $\square$

#### 4.1 An explicit example

Let  $V = \mathbb{A}_{\mathbb{Q}}^1$ , and for each  $t \in V(\mathbb{Q})$  we assign the elliptic curve

$$E_t : y^2 = x^3 + 3(1-t)tx + 2(1-t)^2t.$$

Then  $\Delta(t) = -12^3(1-t)^3t^2$  and  $j(t) = -12^3t$ . The variety  $\Delta(t) = 0$  covers the values of  $t$  for which  $E_t$  is singular. If  $(t_0 : t_1)$  are homogeneous coordinates for  $t$ , then we can look at the homogeneous polynomial  $\Delta(t_0, t_1) = -12^3(t_0 - t_1)^3t_1^2$ , which shows that if  $E_t$  has bad reduction at  $p$ , then  $p = 2, 3$  or  $p|t_1$  or  $t_0 - t_1$ . Using asymptotic estimates for the Euler  $\phi$ -function, we find that

$$|B_K(x)| = \frac{12}{\pi^2}x^2 + O(x \log x).$$

In this case,  $V_{l,p,t} \rightarrow V_{p,t}$  will be finite étale if  $E_t$  has good reduction mod  $p$ , hence we may define  $\Omega_x$  as follows

$$\Omega_x = \{t \in B_K(x) : E_t \text{ is singular, or has bad reduction } \exists p \in \mathcal{L}^*\}$$

The expression of  $\Delta(t)$  above shows that if we have bad reduction at a prime  $p$ , then  $p = 2, 3$  or  $p|t_1$  or  $t_0 - t_1$ . If we try to omit primes  $p$  to get good reduction, the condition  $H(t) \leq x$  tells us we must omit all primes  $p \leq x^{\frac{1}{2}}$ . But the large sieve gives us the best result if we take  $Q \leq x^{\frac{1}{2}}$ , and this causes  $\mathcal{L}^*$  to be empty. On the other hand suppose that  $\mathcal{L}^*$  includes all primes  $p \leq x^{\frac{1}{2}}$ . The  $t_0, t_1$ , are integers with absolute value  $\leq x$ , such that  $t_1, t_0 - t_1$  are not divisible by any prime  $p \leq x^{\frac{1}{2}}$ . In particular, the sieve of Eratosthenes shows that  $|t_1|$  must be a prime in the interval  $(x^{\frac{1}{2}}, x]$ , and so it follows that

$$|C_K(x)| \gg \left( \frac{x - 2x^{\frac{1}{2}}}{\log x} \right)^2.$$

This gives us property 3 with  $\kappa = 2$  and  $\varepsilon = \frac{1}{2}$ , hence in this specific case, the estimate given by (17) becomes

$$\frac{|E_{\mathbb{Q}}(x)|}{|C_{\mathbb{Q}}(x)|} \ll \frac{(\log x)^{8\gamma+3}}{x^{\frac{1}{2}}}.$$

In particular, the map  $j$  is nice enough to allow Duke's original result to be recovered from this.

## **A List of notation**